Ellingson et al. BMC Bioinformatics 2010, 11(Suppl 4):P24
http://www.biomedcentral.com/1471-2105/11/S4/P24

BMC Bioinformatics

POSTER PRESENTATION Open Access

Serendipitous discoveries in microarray analysis

Sally R Ellingson[1*], Charles A Phillips[2], Randy Glenn[3], Douglas Swanson[3], Thomas Ha[3], Daniel Goldowitz[3], Michael A Langston[2]

From UT-ORNL-KBRIN Bioinformatics Summit 2010
Cadiz, KY, USA. 19-21 March 2010

Background

Scientists are capable of performing very large scale gene expression experiments with current microarray technologies. In order to find significance in the expression data, it is common to use clustering algorithms to group genes with similar expression patterns. Clusters will often contain related genes, such as co-regulated genes or genes in the same biological pathway. It is too expensive and time consuming to test all of the relationships found in large scale microarray experiments. There are many bioinformatics tools that can be used to infer the significance of microarray experiments and cluster analysis.

Materials and methods

In this project we review several existing tools and used a combination of them to narrow down the number of significant clusters from a microarray experiment. Microarray data was obtained through the Cerebellar Gene Regulation in Time and Space (Cb GRiTS) database [2]. The data was clustered using *paraclique*, a graph-based clustering algorithm. Each cluster was evaluated using Gene-Set Cohesion Analysis Tool (GCAT) [3], ONTO-Pathway Analysis [4], and Allen Brain Atlas data [1]. The clusters with the lowest p-values in each of the three analysis methods were researched to determine good candidate clusters for further experimental confirmation of gene relationships.

Results and conclusion

While looking for genes important to cerebellar development, we serendipitously came across interesting clusters related to neural diseases. For example, we found two clusters that contain genes known to be associated with Parkinson's disease, Huntington's disease, and Alzheimer's disease pathways. Both clusters scored low in all three analyses and have very similar expression patterns but at different expression levels. Such unexpected discoveries help unlock the real power of high throughput data analysis.

Author details
[1]Genome Science and Technology Program, University of Tennessee, Knoxville, TN 37996, USA. [2]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA. [3]Centre for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, BC V5Z 4H4, Canada.

Published: 23 July 2010

References

1. Allen Brain Atlas: **Home.** *Allen Institute for Brain Science. Web* 2009 [http://www.brain-map.org/].
2. **Cb GRiTS Database.** *Web* 2009 [http://grits.dglab.org].
3. GCAT: **Gene-set Cohesion Analysis Tool.** *The University of Memphis. Web* 2009 [http://binf1.memphis.edu/gcat/].
4. **Intelligent Systems and Bioinformatics Laboratory.** *Web* 2009 [http://vortex.cs.wayne.edu/ontoexpress/].

doi:10.1186/1471-2105-11-S4-P24
Cite this article as: Ellingson *et al.*: Serendipitous discoveries in microarray analysis. *BMC Bioinformatics* 2010 **11**(Suppl 4):P24.

* Correspondence: sellings@utk.edu
[1]Genome Science and Technology Program, University of Tennessee, Knoxville, TN 37996, USA